International Institute for Applied Systems Analysis
IIASA www.iiasa.ac.at

A Stochastic Approach to Model Uncertainty

Fedra, K.

IIASA Working Paper

WP-79-063

August 1979

A STOCHASTIC APPROACH TO MODEL UNCERTAINTY
A Lake Modelling Example

K. Fedra

August 1979
WP-79-63

INTERNATIONAL INSTITUTE FOR APPLIED SYSTEMS ANALYSIS
A-2361 Laxenburg, Austria

PREFACE

Problems of model uncertainty in the Modeling of Environmental Quality Control and Management have attracted recent interest within the Resources and Environment Area, Task 2. Besides Hydrophysical and Ecological Models for Water Quality (including models for the eutrophication process), Uncertainty, Forecasting and Management of Environmental Quality are addressed as a major research topic and incorporated in the Research Plan for 1979-1983 as subtask 2b. Within this framework, this report presents a stochastic approach to the mathematical modelling of uncertain and badly-known systems, using a lake modelling example. The explicit inclusion of data uncertainty in the numerical approach is advocated as a rational means to estimate model output accuracy and credibility. The approach suggests the use of a somewhat fuzzy description of the systems studied in terms of a behaviour space region, taking into account data uncertainty and the stochastic variability of complex natural systems. A corresponding data space region is then established for a model instead of a deterministic data-input vector. Using random samples from this data-input space for simulations, the model output is described in terms of a probabilistic behaviour space. The approach, developed in the context of a simple lake eutrophication model, is suggested for a rather general applicability in the modelling of uncertain natural systems.

ACKNOWLEDGMENTS

This report has been prepared in cooperation with the Austrian Lake Eutrophication Program, Projekt Salzkammergutseen, and the Institute for Limnology, University of Vienna. I most thankfully acknowledge the data and information made available by these institutions and in particular by Dr. Mueller and Dr. Neuhuber.

I also want to thank Dr. Imboden for making available his simulation program SEEMOD2 and for his helpful introduction to the use of the programme.

My special thanks are due to Dr. van Straten and Dr. Beck, who considerably contributed to this study in our weekly discussions.

ABSTRACT

A stochastic approach for modelling uncertain and incompletely known ecosystems, using a lake modelling example, is proposed. In order to estimate the reliability and precision of model predictions based on uncertain data from ecological systems, the explicit inclusion of the uncertainty in the numerical modelling approach is advocated. Starting with a fuzzy definition of systems behaviour in terms of a behaviour space region, the corresponding region in the data space of a given model is explored by Monte Carlo techniques. A set of data vectors--random samples from the data space region corresponding to the empirical range of systems behaviour--is then used to generate independent estimates of states or outputs for selected deterministic inputs. These estimates have to be understood as random samples from a probabilistic behaviour space which reflects the initial uncertainty in data space delimitation. The estimates are used to establish probability distributions for systems states or outputs (cross-sections of the probabilistic behaviour space) for the given input conditions. These probability distributions replace the deterministic point-estimates of a traditional approach, and reflect the incomplete knowledge about the system as well as the stochastic variability of ecosystems. The approach is extended for long-term simulations of systems behaviour under changed input conditions, and estimates of prediction accuracy in time are obtained.

CONTENTS

A STOCHASTIC APPROACH TO MODEL UNCERTAINTY
A Lake Modelling Example

K. Fedra

INTRODUCTION

Uncertainty and Arbitrariness in Ecosystems Models

Modelling of ecological systems is certainly an important tool of the "hard science" of systems ecology (Patten 1971). However, there are many elements of "soft science", namely, uncertainty, arbitrariness, and chance, which--although computerized--severely affect the credibility of predictions from mathematical models of ecological systems.

Mathematical models of ecosystems are--more less necessarily as a consequence of the homomorphic modelling approach--drastic simplifications. As a rule they do not allow for various typical features of ecological systems as well as of data describing such systems. With regard to the systems themselves, these features would be richness and variety, spatial heterogeneity, nonlinearity, functional dissimilarity within lumped components, and stochastic variability. With regard to the data a modeller has usually at hand for the analysis of a given ecosystem, these data are--at least in terms of a deterministic mathematical model--scarce, scattered and uncertain, and often enough inadequate in light of the posed problems and the desired accuracy in their solution. However, having in mind as one ultimate goal the

application of mathematical models as a rational tool for ecosystems management and control, the uncertainty of models and their relative precision must be evaluated under the above constraints. Much of the uncertainty in ecosystem modelling seems to be an essential part of the objects studied, or it is simply a scale and effort problem as in environmental data collection. And much of the arbitrariness--just consider standard sampling strategies in the time and space domain, the criteria for selecting one specific model, or the choice of objective functions in calibration methods--seems unavoidable. Often there is simply not enough information available for more rational decisions. However, all these sources of uncertainty and arbitrariness affect model predictions. It is therefore an important task to make explicit the effects of arbitrary assumptions and uncertainties in our knowledge. Being aware of the shortcomings and insufficiencies of environmental data and modelling techniques, by explicitly including the uncertainty in our analysis of ecosystems, we should at least be able to estimate the level of accessible precision of predictions.

Taking advantage of modern computer technology, a straightforward trial-and-error approach has been chosen in favor of more sophiscated analytical methods. If arbitrary assumptions have to be made at all, why not make several of them equally good in terms of our incomplete knowledge, and investigate the whole range of this arbitrariness in its effect on the results. If we know that the data we have to use are uncertain, why not explore the whole range of uncertainty in its relation to the credibility of the output of our analyses. Having a particular environmental problem to solve and having in hand a mathematical model* which is supposed to be appropriate for that purpose--which

---

*The author is well aware of the fact that already the selection of any one model includes a first element of arbitrariness; it seems most likely, that different models, although using the same set of data, will give somewhat different results. However, this source of uncertainty, which clearly introduces an additional dimension of the problem, is not considered here.

means a so-called validated model--we find ourselves confronted with the problem which numbers to put into the model to get the desired answers. The desired answers are usually of the kind: what will the systems' behaviour be under such and such input conditions (in the future, of course)? To answer such questions by means of a mathematical model, certain data are required. The data-input requirement of a dynamic, non-homogeneous model could conveniently be grouped in the model parameters sensu stricto, import and forcing describing data, and the initial conditions. To estimate these values, we have to use the observed (past) system behaviour as well as experimental evidence and information from the literature. However, all these data (as a field-ecologist maybe recognizes rather than a modeller) are only rough estimates of systems properties (which are described in the model on a high level of aggregation), connected with uncertainties resulting from the above listed peculiarities of ecological systems as well as the limited possibilities of ecological field studies. It is obvious that the use of any such uncertain data will cause consequent uncertainty in the model output, not to speak of other principle problems in model predictions (see p. 4). For recent approaches to consider and include aspects of uncertainty, alternative to traditional deterministic modeling techniques, see Spear and Hornberger (1978), Di Toro and van Straten (1979) Beck et al., (in press).

An important method to derive the model data requirement from the available information on the system studied is the calibration of model parameters. Traditional parameter calibration methods (e.g., Lewis & Nir 1978) refer to a short and well defined period of the systems' history. The inputs and forcings as well as the initial conditions for this period are assumed to be exactly known. According to the objective function chosen (as a rule least square approximation of the available data "points" from the time series), one "optimum" (by definition) solution is obtained. However, the "points" from the time series must be considered as ranges, and the parameter vector obtained is just one out of a set of parameter vectors or a parameter space of generally unknown extension.

And the relationship of parameter space extension (introducing a more general concept of model data-input space extension) and model prediction accuracy (the behaviour space extension for a given "input" situation of some uncertainty) will be the specific topic of this investigation. The final goal will be to show that the prediction of systems states and outputs, using uncertain data-input by necessity, has to be reformulated in terms of probability distributions. These probability distributions represent the extent of uncertainty resulting from our incomplete knowledge of the systems previous states (which we use to estimate model parameters and initial conditions) as well as the natural stochastic variability in the future imports and forcings. But even this fairly advanced concept is based on several simplifying assumptions, which further add to the prediction uncertainty. The approach does not consider uncertainty in the model structure: it assumes (as a technically necessary simplification) that the model chosen is appropriately representing the system studied over the whole range of input conditions. Another of these assumptions is that parameters, estimated from a certain range of input conditions (or a certain region in state space) will also be valid outside this range, that they are state- and input-independent. This would imply that systems do not adapt to changes in their environment, that they do not change their structure as well as the rates of their processes under different conditions--which is most obviously not true (Straskraba 1976, Fedra 1979).

In order to achieve probability distributions for system states under uncertain input conditions, some straightforward methods of simulation and analysis are proposed. To summarize the approach (see also Figs. 1 and 3), it first describes the behaviour of a system for a sufficiently long period (several years, if possible), for which, however, the system must be assumed to be in some kind of dynamic "steady state" or oscillating within certain limits. The description of behaviour includes the stochastic variability of ecological systems in time as well as the uncertainty in the underlying observations. Therefore, ranges are specified for a set BR of m behaviour

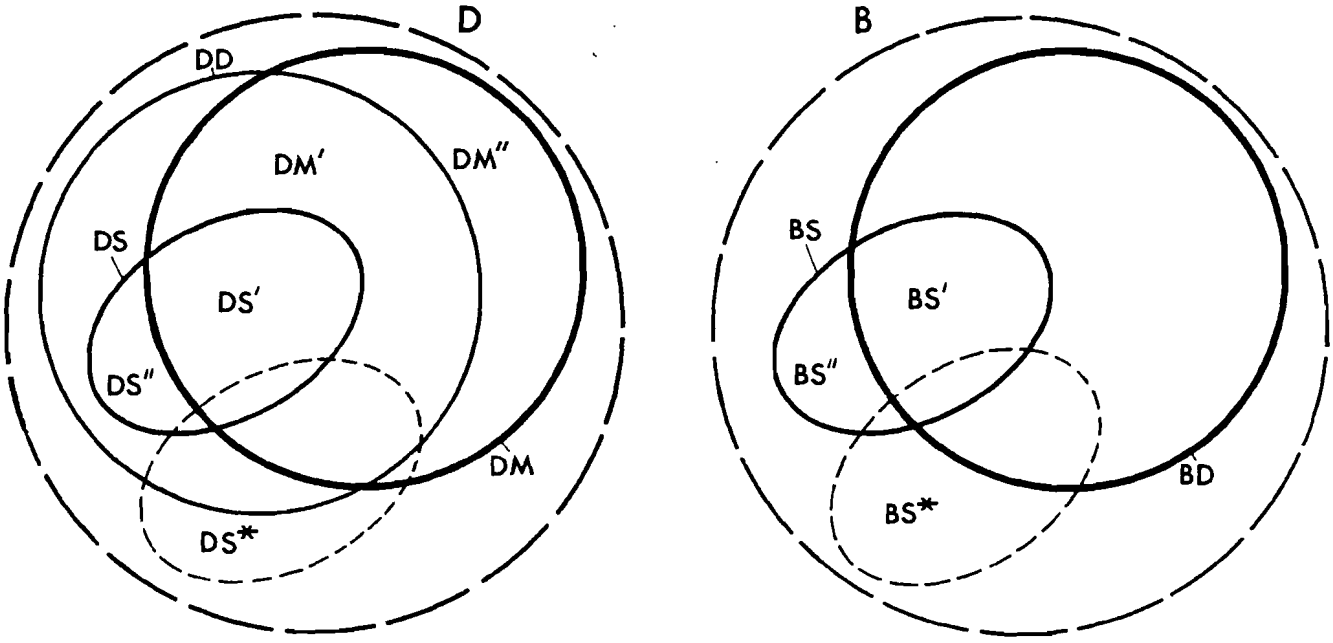

Figure 1. Data input vector set D and Behaviour set B, showing the relations of their respective subsets. For further explanation see text.

describing measures BR_i (by defining $BRMIN_i$ and $BRMAX_i$), which are formulated in terms of the simulation model used:

$$BR = \{BR_1, BR_2, \ldots, BR_m\} \quad \text{or}$$
$$BR_i \in BR \quad (i=1,\ldots,m) \quad ,$$

where

$$Br_i = \{b_i \mid (BRMIN_i \leq b_i \leq BRMAX_i)\} \quad .$$

The set of ranges of the behaviour measures BR_i now defines a set BD of allowable behaviours as a proper subset from the set of all possible behaviours B or a region BD in the m-dimensional behaviour space B :

$$BD \subset B \quad ,$$
$$BD = \{BD_j \mid (BD_j = \{b_i (i=1,\ldots,m) \mid (b_i \in BR_i)\})\} \quad .$$

Second, a data space D is defined for the simulation model to be used. Its n dimensions are the data required for the model, namely, the model parameters sensu stricto, import and forcing describing data and the initial conditions. Ample ranges DR_i are established (specifying $DRMIN_i$ and $DRMAX_i$) for each of these data. The ranges DR_i are based on our knowledge about the system, experimental evidence, or the literature, and define a region DD (the set of allowable data vectors) as a proper subset of all possible data vectors D :

$$DR = \{DR_1, DR_2, \ldots, DR_n\} \quad ,$$
$$DR_i = \{d_i \mid (DRMIN_i \leq d_i \leq DRMAX_i)\} \quad ,$$
$$DD = \{DD_j \mid (DD_j = \{d_i (i=1,\ldots,n) \mid (d_i \in DR_i)\})\} \quad ,$$
$$DD \subset D \quad .$$

Third, this data space region DD is now randomly sampled N times by Monte Carlo methods. Each sample data vector DS_i $(i=1,\ldots,N)$ is then used for a simulation run, and the resulting set BS of

of model behaviours BS_i (i=1,...,N) is classified according to the system's behaviour definition BD:

$$BS' = \{BS_i \mid BS_i \in BS) \wedge (BS_i \in BD)\} \quad n(BS') = M \quad ,$$
$$BS'' = \{BS_i \mid (BS_i \in BS) \wedge (BS_i \notin BD)\} \quad n(BS'') = N-M \quad .$$

The set DS of sample data vectors DS_i is accordingly separated into two complementary subsets DS' and DS'', using the relationship given by the model

$$BS_i = f(DS_i) \quad ,$$

such that

$$DS' = \{DS_i \mid (DS_i \in DS) \quad \forall DS_i \; (BS_i = f(DS_i)) \wedge (BS_i \in BD)\} \quad ,$$
$$DS'' = \{DS_i \mid (DS_i \in DS) \quad \forall DS_i \; (BS_i = f(DS_i)) \wedge (BS_i \notin BD)\} \quad .$$

With M and N-M elements, respectively, The set DS' of data vectors DS'_i giving rise to a behaviour BS'_i completely within the definition boundaries are considered as random samples from the data space region DM corresponding to the behaviour definition space region BD of the model:

$$DS' \subset DM \quad ,$$
$$DM = \{D_i \mid (D_i \in D) \; \forall D_i \; (B_i = f(D_i)) \wedge (B_i \in BD)\} \quad .$$

The total sample of M behaviour generating data vectors is then analysed to give some insight into data interdependencies and the data space configuration.

Fourth, the set DS' of M sample data vectors DS'_i

$$DS' = \begin{bmatrix} d_{11},\ldots,d_{1M} \\ d_{21},\ldots,d_{2M} \\ d_{i1},d_{ij},d_{iM} \\ \ldots,\ldots,\ldots \\ d_{n1},\ldots,d_{nM} \end{bmatrix}$$

is then used for further simulations. Any of the elements or combination of L elements (corresponding to the rows in the matrix notation) d_i (i = 1,...,n) can now be varied systematically by substituting

$$d_{ij} = X_k \quad (j = 1,\ldots,M \quad ; \quad k = 1,\ldots,K)$$

to study the model behaviour reactions. For each value X_k (k = 1,...,K) of the systemically varied data vector element d_i a sample of M estimates of behaviour describing measures bs^*_i (i = 1,...,m) is obtained using the M x K data input vector matrix DS*:

$$DS^* \begin{bmatrix} DS^*_{11},\ldots\ldots,DS^*_{1M} \\ DS^*_{21},\ldots\ldots,DS^*_{2M} \\ DS^*_{i1},DS^*_{ij},DS^*_{iM} \\ \ldots\ldots,\ldots\ldots,\ldots\ldots \\ DS^*_{K1},\ldots\ldots,DS^*_{KM} \end{bmatrix} .$$

These samples from the behaviour space are used to establish a m x K matrix of probability distributions PDF_{ij} (i = 1,...,m ; j = 1,...,K) or a set PDF for the subsets BS^*_j of the behaviour space region BS* of the model:

$$PDF_{ij} \in PDF \quad (i = 1,\ldots,m \;;\quad j = 1,\ldots,K) \quad ,$$

$$BS^*_j \subset BS^* \quad (j = 1,\ldots,K) \quad .$$

The set PDF of probability distributions now describes in terms of the model behaviour the effects of the initial uncertainty in the model data input, represented in the variability of the (n-L) data values through the M sets DS^*_j from the data input matrix DS*.

As a concrete example, subjected to such an approach, an analysis of the eutrophication process of an Austrian lake was performed. However, it is intended to show that the principal conclusions of this study also hold true for any complex and uncertain system, subjected to deterministic mathematical modelling.

THE LAKE SYSTEM AND THE ENVIRONMENTAL PROBLEM

In close cooperation with the Austrian Lake Eutrophication Program, Project Salzkammergutseen, the Attersee, a deep, stratified, oligotrophic lake of almost 4000 million cubic metres and a theoretical fill-in time of seven to eight years, was subjected to our approach. Basic lake data are compiled in Table 1. The investigations on the lake, carried out since 1974 within the frame of the OECD Lake Eutrophication programme, indicated a slow but steady eutrophication trend for the lake. As a main source of phosphorus loading, the discharge of the upstream Mondsee, a smaller (510 Mill.m^3) but more eutrophic lake, was identified as the major source for eutrophication.

However, the problem setting is somewhat diffuse. The major items in terms of a possible management application are missing: there are no well defined objectives (maintaining "sufficiently good" water quality?), nor constraints, nor alternatives, which could be stated in quantitative terms. This is partly due to the multiparametric nature of the "eutrophication problem", which should rather be called a "diffuse concern". It was therefore necessary to operationally define one (of course debatable) measure for eutrophication or water quality, and select one or a

Table 1. Attersee basic lake data.

Surface area:	45.9	km^2
Catchment area:	463.5	km^2
Maximum depth:	171	m
Mean depth:	84	m
Volume:	$3.9*10^9$	m^3
Length:	20	km
Average width:	2	km
Theoretical fill-in time:	7 - 8	years
Average hydraulic loading:	17.5	m^3*sec^{-1}

After Floegl 1974

few key issues for further study. In light of the above limitations, it is obvious that the goals of the study in environmental terms are not so much management advice but rather insight in selected key processes and the relationship of accessible prediction accuracy to various sources of uncertainty in modelling the Attersee.

The question selected for our modelling approach was the relation of the trophic state of the Attersee (measured as yearly primary production per unit lake area) to the import of particulate as well as dissolved phosphorus. The proportion of the Mondsee discharge in the total phosphorus loading of the Attersee was estimated with about 50 to 60%. However, absolute as well as relative estimates for the loadings have "confidence limits" of approximately +/- 50%. Similar uncertainties are found in the estimates of lake phosphorus concentrations. Therefore, the available time series of the phosphorus data could not be used for a traditional parameter calibration method. This is partly also due to the fact that the orthophosphate level in Attersee is usually around 1 $mg*m^{-3}$, which corresponds to the level of detectability. These problems were aggravated by the spatial, vertical as well as horizontal heterogeneities in the lake: due to the shape and the inflow in the southernmost part of the channel-like lake, a south-north gradient in the nutrient levels was assumed. However, measurements from two stations,

one situated in the south, close to the inflow, and the other one in the northern part of the lake, showed no significantly different phosphorus level. For simplicity (and with regard to the lack of data supporting a more sophisticated physical framework), the lake was therefore considered as a horizontally completely mixed water body.

METHODS

The definition of a "typical system behaviour" is a crucial concept in this approach and warrants some considerations: traditional parameter calibration methods search for one single point in parameter space, usually through the optimization of some objective function, referring to observed "system behaviour". Systems behaviour in this context means a deterministic trajectory in one or more state dimensions. Even if these trajectories are referred to as ranges (allowing for measurement uncertainty, see Di Toro and van Straten 1979), it is the dynamic behaviour (a set of tuplets: time-value, as a rule of one single year) of the system which is referred to. However, much of the information we have about ecological systems is non-dynamic in nature; and much of this non-dynamic information may be comparatively certain, due to long term integration or the pooling of measurements in aggregated values. The selection of one specific year of system behaviour or one specific day chosen for a measurement of course contains some arbitrariness, which is at least less when we use long term averages, cumulative or integrated properties instead (provided such data are available!). Pooling of data for such measures does not mean a loss of reliable information at all: whereas single measurements contain much (generally unknown) uncertainty, the pooling of measurements allows us to estimate the reliability of a derived estimate in terms of confidence limits. As the proposed approach uses parameter space delimitation instead of parameter calibration, and therefore a simple "objective function", non-dynamic information can easily be used. Each of the available measures to describe the typical (long term) system behaviour defines one dimension in state or output space, and the ranges for each

of these measures define a multidimensional box in what I would like to call behaviour space. The only thing the "unintelligent calibration routine" then does is to check whether a randomly chosen point in data space gives rise to a model behaviour which completely lies within the behaviour space box or not.

The behaviour definition used has to include enough of the available information to allow for a typical and realistic picture of the Attersee behaviour in terms of the simulation model output. Seven measures were selected, based on the available data and in discussion with the involved biologists. Ranges were then specified for each measure. This takes into account the incomplete knowledge about the system (measurement uncertainty) as well as the natural stochastic variability of the ecosystem, most obvious in the year-to-year differences in certain system states. The measures applied and their allowable ranges are:

1. Total primary production has to be between 50 and 150 $gC*m^{-2}$ and year.
2. Total phosphorus export per year has to be between 2 and 8 tons.
3. The peak value of particulate phosphorus in the epilimnion has to occur between day 60 and day 210 (relative to January 1, start time of the simulations).
4. The peak value of particulate phosphorus in the epilimnion must not exceed 15 mg $P*m^{-3}$.
5. The concentration of phosphate during the mixed period must not exceed 2.5 mg $P*m^{-3}$.
6. The peak value of particulate phosphorus must at least be two times the minimum value.
7. The maximum total phosphorus content of the lake during the year must not exceed two times the minimum value.

These definitions of the Attersee behaviour in terms of model output data can now be understood as a seven-dimensional box in behaviour space, within which the model behaviour for a correspondingly defined empirical range of "input" and initial conditions (see Table 2, data 11-13, 17-22) has to be.

DATA SPACE DELIMITATION: A CALIBRATION ALTERNATIVE

The term data vector as it is used throughout this paper, subsums the total data requirement of a mathematical simulation model. This includes, besides the model parameters sensu stricto, the initial conditions, and forcing- and import-describing data. For an explanation of the description of time-varying parameters and forcings by means of these data (d11-d20) see p. 16, Figure 2.

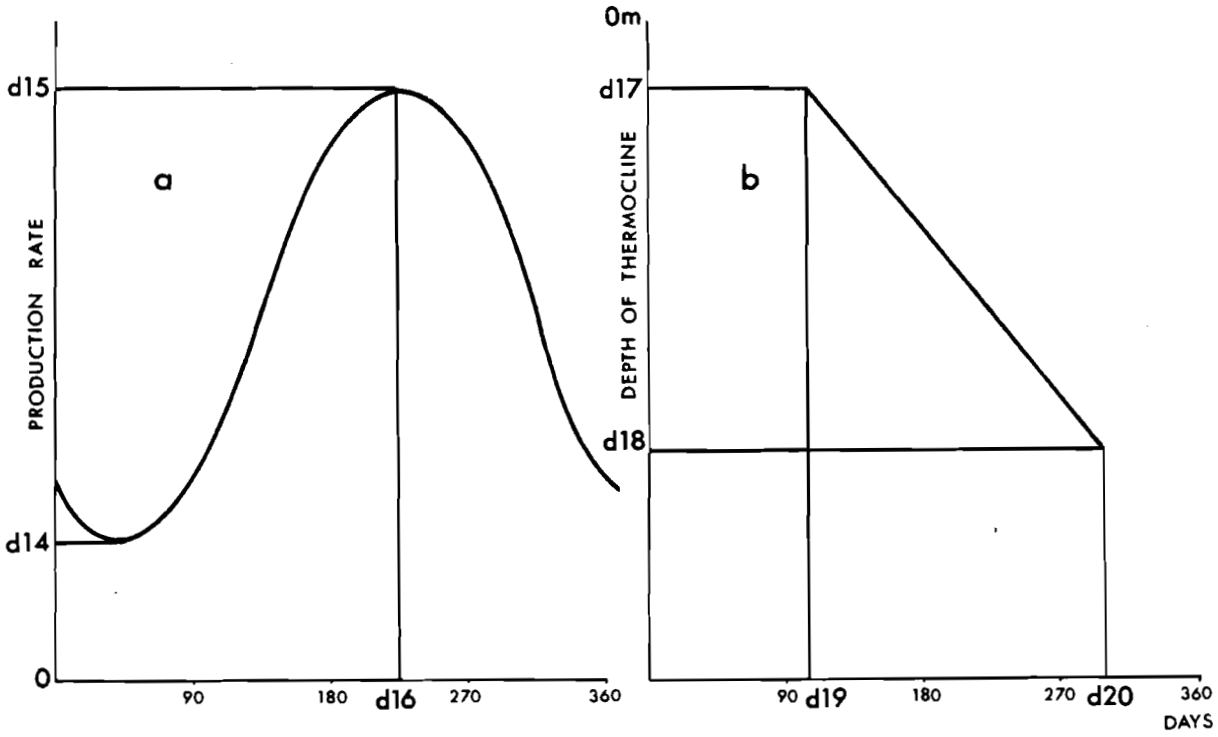

Figure 2. Description of time-varying forcings by constant parameters: (a) particulate phosphorus production rate: d14 = minimum production rate; d15 = maximum production rate; d16 = time lag of maximum relative to start (January 1). (b) depth of thermocline: d17 = initial thermocline depth; d18 = final thermocline depth; d19 = begin of stratified period; d20 = end of stratified period.

Table 2. Initial data-input space definition

DATA TYPE	UNIT	MINIMUM	MAXIMUM
Parameters sensu stricto:			
1 Michaelis constant (phosphorus)	mg^{*-3}	0.20	2.00
2 resp./mineralization epilimnion	day^{-1}	0.02	0.20
3 resp./mineralization hypolimnion	day^{-1}	0.01	0.05
4 net sedimentation velocity epilimnion	$m*day^{-1}$	0.01	1.00
5 net sedimentation velocity hypolimnion	$m*day^{-1}$	0.02	2.00
6 diffusion coefficient hypolimnion	cm^2*sec^{-1}	0.02	0.50
7 diffusion coefficient thermocline	cm^2*sec^{-1}	0.01	0.25
8 extinction coefficient	m^{-1}	0.20	0.40
9 self shading coefficient	m^2*mg^{-1}	0.01	0.02
10 thickness of thermocline	m	5.00	10.00
Import- and forcing describing data:			
11 orthophosphate import	$mg*m^{-2}*day^{-1}$	0.01	0.20
12 particulate phosphorus import	$mg*m^{-2}*day^{-1}$	0.25	1.50
13 hydraulic loading	$m*day^{-1}$	0.02	0.05
14 minimum production rate	day^{-1}	0.25	0.50
15 maximum production rate	day^{-1}	1.00	10.00
16 time lag of production maximum	day	180	270
17 initial thermocline depth	m	3.00	6.00
18 final thermocline depth	m	15.00	20.00
19 begin of stratified period	day	120	280
20 end of stratified period	day	280	330
Initial conditions:			
21 initial orthophosphate mixed period	$mg*m^{-3}$	0.20	2.00
22 initial particulate P mixed period	$mg*m^{-3}$	2.50	7.50

Only a few values, assumed to be certain such as lake volume, surface area or depth--which one could call site variables--are excluded. Altogether, the model used in this study (see p. 16) required 22 data values. Based on the available field data (ATTERSEE: Vorlaeufige Ergebnisse des OECD Seeneutrophierungsund des MaB-Programms 1976, 1978) as well as on the literature, ranges were established for each of these data. Thus a region in the 22-dimensional data space was defined.

The ranges were chosen such that for measures, which are directly physically interpretable--e.g., the input describing coefficients for ortho- and particulate phosphorus loading and the hydraulic loading--the available estimates were used as the mean. The ranges were then extended symmetrically according to the observed variability of the respective measures. More complex and not directly interpretable parameters (e.g., the algae growth rate or the respiration/mineralization rate) were given ample ranges with means approximated to values from the literature.

Although there is some ambiguity in the classification of the data types, a separation in traditional groups is attempted in the table above.

Given the definition of the range of typical systems behaviour in terms of the model output (p. 12), the initial data space should separate in regions giving rise to that behaviour and regions not giving rise to that behaviour. The data space for the model and thus (partly) corresponds to a parameter calibration method. As there was no a priori information about regions of special interest, where the search for appropriate data vectors could be concentrated by assuming specific probability distributions within the ranges for the individual parameters and data, simple rectangular probability density functions were used.

THE NUMERICAL APPROACH

In order to study the loading/production relationships of Attersee, an appropriate available simulation model was selected: the model SEEMOD2 (see Imboden and Gaechter 1978), which predicts primary production per unit lake area as related to import of soluble reactive as well as particulate phosphorus and various forcings and model parameters, was chosen. The relationship between phosphorus loading and primary production is described by means of a dynamic, one-dimensional, vertical diffusion model for soluble reactive phosphorus and particulate phosphorus. It takes into account Michaelis Menten kinetics and self shadowing of algae. The model uses the lake morphometry, hydraulic loading, respiration, sedimentation, vertical eddy diffusion, and depth of thermocline.

Some minor modifications of the model were made in order to allow the description of time-varying forcings (production rate and depth of thermocline) by means of constant coefficients (p. 14, Figure 2). The model originally uses time-varying parameters and forcings which are read from tables in the form of tuplets time-value. Actual values for each integration step are determined by linear interpolation. In order to reduce the number of data required, auxiliary coefficients were defined, and the time-varying values are described by means of these coefficients as functions of time.

For example, the time-varying production rate mu is described as follows:

```
ZEIT = TIME - 81. - d16
MU = SIN(3.1416/180.*ZEIT)
MU = MU*(d15-d14)/2. + (d15-d14)/2.
```

where TIME is the current simulation time (days), d14 and d15 give the minimum and maximum for the production rate, and d16 is the time lag of the maximum relative to January 1, the starting day of the simulations (see Figure 2). Other potentially time-varying data were kept constant, as the available field data did not allow the specification of a meaningful pattern.

With regard to the morphometric situation of Attersee, the optional backflux of phosphorus from the sediments was set to zero. Altogether, the data requirement of the model (excluding "known" parameters such as surface area, maximum depth, etc.) included 22 "parameters" (see p. 13). The simulation model SEEMOD is a comparatively simple model--in the array of available lake models--especially in its biological aspects. The model does not describe zooplankton explicitly. The effects of zooplankton on the algae are included in the first-order loss term respiration/mineralization. The purpose of the model must consequently be seen in the prediction of loading/production relationships on a yearly time scale rather than in the prediction of short term algae/phosphorus dynamics.

The simulation model SEEMOD was incorporated as a subroutine in a control programme MONTEC, which randomly generated data-input vectors out of the initial data space, started the simulation model, and saved the relevant output (see Figure 3, cycle 1). Each run required approximately 10 CPU sec and 18 k bytes (PDP11/70 under UNIX). Output of each run consisted of 52 floating point words, and included run number, random generator seeds, the 22 elements of the input vector, end time of the simulation run (as a check for aborted simulations) and 26 model output data.

Altogether, 1000 runs of the model, using the initial data-input space, were performed in cycle 1. The output of these runs was subjected to the analysation programme ANALYSE1, which separated the output in two groups, GOOD and BAD, according to the behaviour definitions listed in Table 2. The programme ANALYSE1 determined the ranges, means, and standard deviations of the data and certain output values for the two groups of runs. To investigate the influence of the behaviour definition itself on the data-input space seggregation, the programme ANALYSE1 was also run with the definition ranges and boundaries confined and extended for 10% of their initial values. Plots were then made of the frequency distributions of the parameters for the behaviour class. They were established using 10 classes of 1/10 of the originally estimated ranges for each of the data.

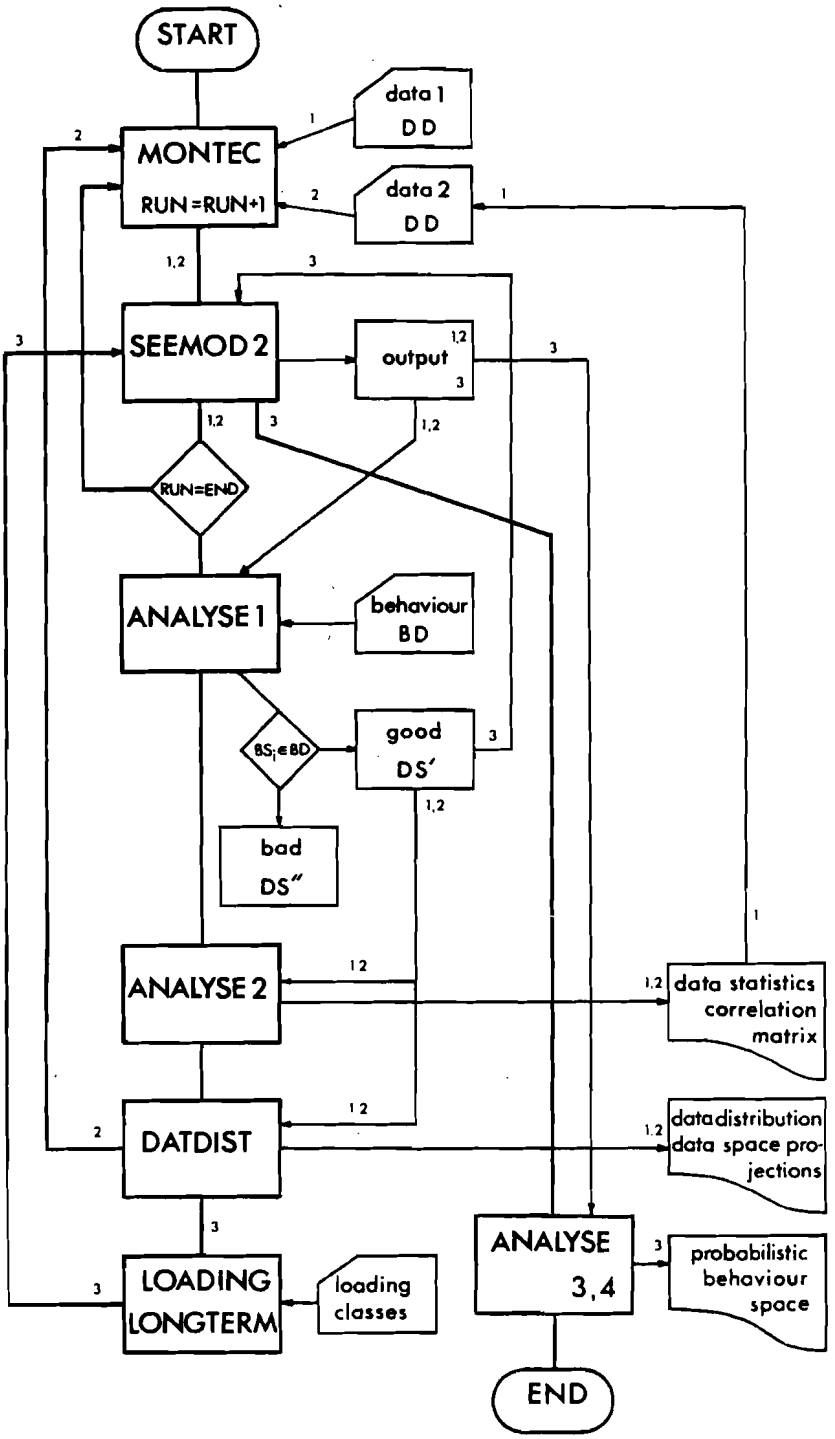

Figure 3. Flow chart of the numerical approach. Capital letters denote programmes, lower case for data files; thick lines indicate programme sequence and control flow; thin lines represent I/O procedures; numbers designate subsequent cycles.

The two behavioural classes GOOD and BAD (including a few aborted runs) were then analysed using the programme ANALYSE2, which established in addition to the basic statistics a parameter/output correlation matrix. Based on the results of this analysis, new confined ranges were established for 5 out of the 22 data. With this new parameter space definition another 10,000 runs were made, resulting in 293 behaviour generating vectors. The same analyzation programmes were used as for the first series of simulations in order to study the data-input space configuration. Since the 293 behaviour vectors are independent random samples from the data space corresponding to the defined behaviour space, they were now used for the subsequent analysis of loading/production relations and for the study of long-term system reactions to different loadings. A summary of the numerical approach is presented as a flow chart in Figure 3.

RESULTS

Data Space Configuration

Only 56 out of the 10,000 sample data vectors from the initial ranges were giving rise to a model output completely within the behaviour space region corresponding to the behaviour definition (Figure 3, cycle 1). A comparison of data statistics for the two classes GOOD and BAD shows that the standard deviations from the behaviour class are always lower than the corresponding values from the non-behaviour class. Ranges and means, however, showed no obvious differences in most cases. This was partly to be expected--for the means at least--due to the specification of the initial data ranges. The ranges for five out of the 22 data could be confined on the basis of this first 10,000 runs. The resulting second set of data ranges was now again randomly sampled 10,000 times (Figure 3, cycle 2). Two hundred and ninety-three data vectors from this cycle (Table 3) were found to generate a model behaviour according to the definition (p. 12).

Table 3. Statistics of the 293 behaviour generating data sets.

DATA		mean	S.D.	minimum	maximum
1	Michaelis constant	1.00	0.51	0.20	1.99
2	Respiration (epilimnion)	0.11	0.05	0.02	0.02
3	Respiration (hypolimnion)	0.014	0.003	0.010	0.024
4	Sedimentation (epil.)	0.26	0.16	0.01	0.74
5	Sedimentation (hypol.)	1.42	0.42	0.32	2.00
6	Diffusion (hypolimnion)	0.28	0.14	0.02	0.50
7	Diffusion (thermocline)	0.13	0.07	0.01	0.25
8	Extinction coeff.	0.29	0.06	0.20	0.40
9	Self shading coeff.	0.015	0.003	0.010	0.020
10	Thickness of thermocline	7.54	1.47	5.01	9.98
11	OP import	0.107	0.051	0.012	0.199
12	PP import	0.929	0.334	0.260	1.500
13	Hydraulic loading	0.042	0.005	0.030	0.050
14	Production rate minimum	0.377	0.073	0.251	0.499
15	Production rate maximum	6.343	2.230	1.180	10.000
16	Time lag of d15	218.5	25.4	180.0	269.8
17	Initial thermocline depth	4.47	0.85	3.01	5.99
18	Final thermocline depth	17.64	1.41	15.03	19.97
19	Begin of stratified period	155.4	16.8	120.9	179.9
20	End of stratified period	302.6	14.9	280.0	329.8
21	Initial OP	1.065	0.493	0.200	1.989
22	Initial PP	3.445	0.713	2.516	6.107

To investigate the effect of the behaviour definition itself on the data space separation, ANALYSE1 was also run with the behaviour definition boundaries confined and extended for 10% of their initial ranges, using the first set of 10,000 data vectors from cycle 1. As expected, the number of behaviour runs was fewer with the narrow definition set (5 behaviour runs as compared to 56 with the original definitions) and larger with the wide definition set (203 behaviour runs).

However, the data means for the three groups (narrow, original, wide) as well as the output values investigated did not differ significantly. This indicates, that the arbitrariness in the choice of the definition boundaries does not influence the centre of gravity of the behaviour generating data space region. This assumption was substantiated by plots of the frequency distributions of the data values from the behaviour group (10 classes over the whole initial range, see Figure 4). Most of them showed clearly centralised distributions, where the kurtosis of the frequency distribution could be used as a measure of model sensitivity (under the behavioural constraints) for the respective data. Another indication of increasing "density" towards the centre of gravity was found using a third "narrow" data space definition. Data ranges were reduced to 60-80% of their initial extent, centred around the means of the first 17 behaviour cases obtained during the first series of runs. This narrow data space definition gave a score of behaviour vectors of 15% as compared to 0.6% for the initial range and 3% for the second data range definition. The distribution and largely overlapping ranges of data from the two behaviour classes make obvious that one and the same value for certain data could give rise to the behaviour or not, depending on the other elements of the data vector. It is therefore the data combinations that influence the model behaviour rather than the absolute values of the single data--at least within a certain range of the values. The programme ANALYSE2 was therefore used to establish a data correlation matrix for the behaviour classes (Table 4).

As expected with regard to the random data vector generation, there was no significant interdata correlation in the BAD class. In comparison, the data of the behavioural class showed distinct interdependencies. Significant positive correlation was found, e.g., between sedimentation velocity in the

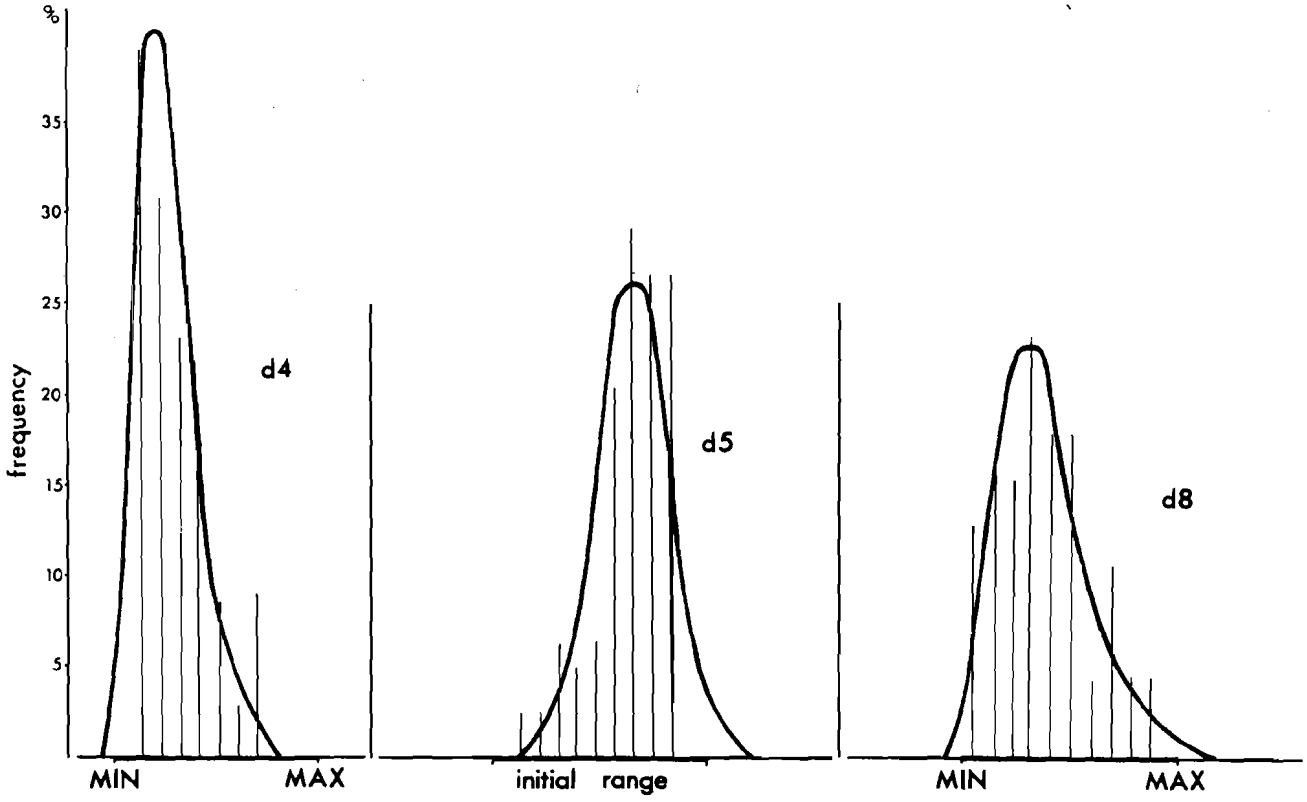

Figure 4. Input data frequency distributions for the behaviour class GOOD cycle 1. Class width is 1/10 of the initial data range; curves show probability distributions fitted for the frequency data. d4: net sedimentation velocity epilimnion; d5: net sedimentation velocity hypolimnion; d8: time lag of production maximum.

Table 4. Data-input correlation matrix.

	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22
1	X														+							
2		X																				
3			X									-	+		+							*-
4				X	-							*+	+			*-					*-	
5				-	X							+	-			+						+
6						X																
7							X					+										
8								X														
9									X													
10										X												
11											X											
12			-	*+	+		+					X	-									
13			+	+	-							-	X									-
14														X								
15	+		+												X							
16				*-	+											X						
17																	X					
18																		X				
19																			X	+		
20																			+	X		
21				*+																	X	
22			*-		+							-										X

Significant correlation ($r > 0.15$; $n = 293$) indicated by "+" or "-", respectively, highly significant correlation ($r > 0.25$) by "*".

epilimnion and begin of stratified period, diffusion coefficient in the thermocline and final thermocline depth, sedimentation velocity hypolimnion and initial particulate phosphorus concentration, sedimentation velocity epilimnion and particulate P import, etc.

Significant negative correlations were found, e.g., between maximum production rate and sedimentation velocity hypolimnion, diffusion coefficient thermocline and sedimentation velocity epilimnion, mineralization hypolimnion and orthophosphate import, etc. However, only 13 out of 22 data showed significant interdependencies. d3 (respiration/mineralization hypol.), d4 and d5 (sedimentation epi- and hypolimnion), d12 (particulate P import), and d13 (hydraulic loading) showed the most complex relations with four or five "dependencies". Also, production determining d15 and d16 as well as the initial conditions d21 and d22 show more than one and partly highly significant dependencies.

The correlation structure can be explained with reference to the function of the individual data in relation to the behaviour defining output values. Data influencing a specific output, used for the behaviour definition, in the same direction and amplifying their effects are found negatively correlated, whereas data with opposed effects are positively correlated. In addition, the correlation matrix can also be read as a table of data-combination sensitivities. These correlations, together with the extent of the data ranges and the distribution of data within these ranges, determine the "shape" of the 22-dimensional data space of the model.

To gain some more insight in the obviously complex and bizarre shape of the data space, plots were made projecting on two-data planes. Using a 50 x 50 grid, plots were made for the 231 combinations from the data matrix. A few examples of such plots are shown in Figure 5. The 56 (293 respectively) points, representing the behaviour vectors, were found to show significant departures from a random distribution pattern in most cases. Most plots show distinct aggregation in certain regions of the projection planes. The high dimensionality of the problem would require some more sophisticated methods for the analysis of the data space. Questions such as whether there exists a unique data

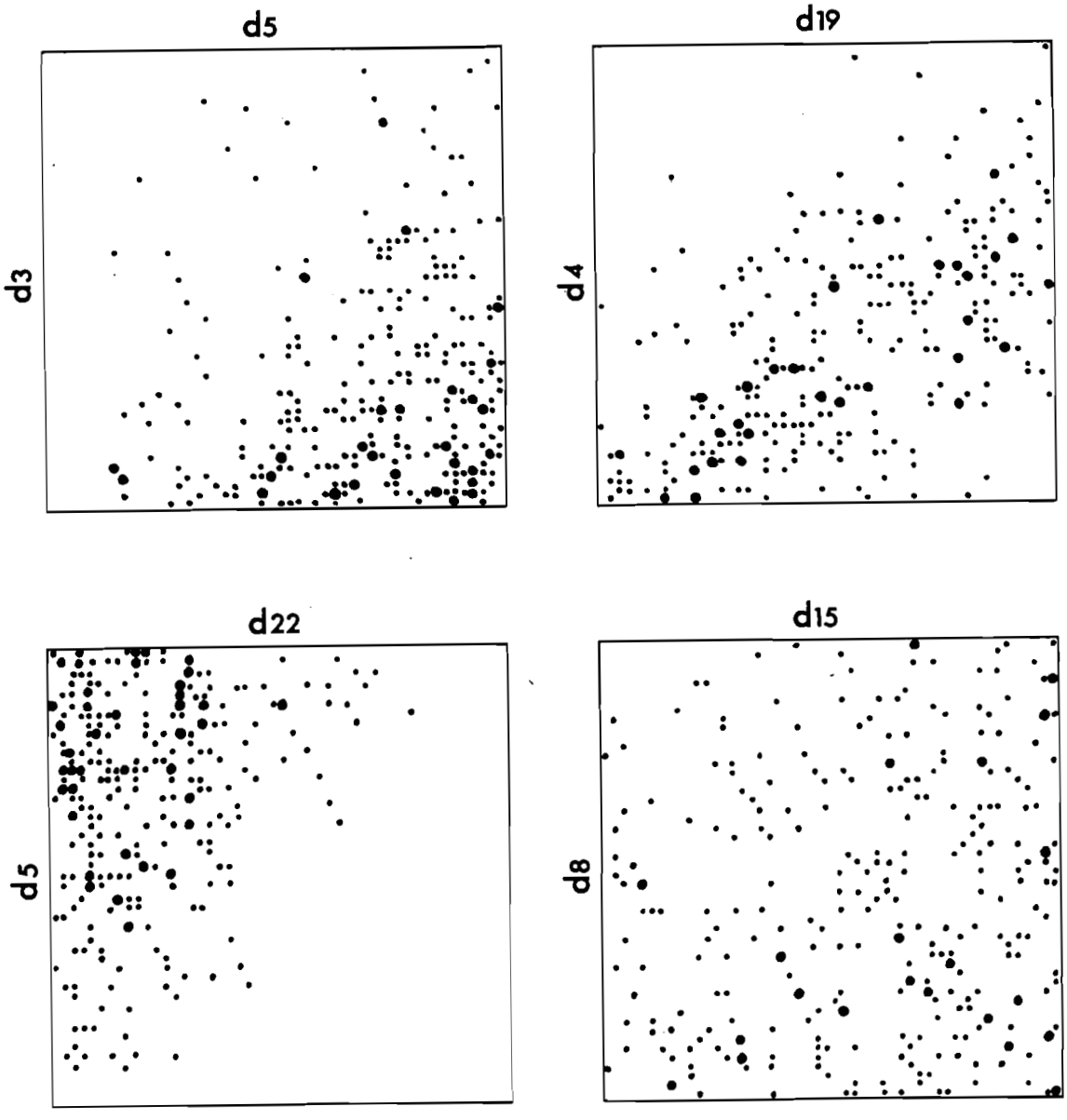

Figure 5. Data space projections on two-data planes. 50 x 50 grid, 293 data vectors from cycle 2. Single vectors indicated by ".", more than one per grid cell by "●".

space region corresponding to a given region in state space, or whether such a region is closed or not, are of course of considerable theoretical interest. However, within the frame and under the limitations of the study presented here, such questions have to be postponed.

THE PRAGMATIC PREDICTION METHOD

In order to study the relationship of phosphorus loading and primary production, the 293 behaviour generating data vectors, independently generated and thus random samples from the "true" but structurally only roughly known data space, were used for another series of simulations (Figure 3, cycle 3). In this series of simulations (each for a one-year's period only) the load-determining coefficients d18 and d19 were now systematically varied from a total loading (d18+d19) of 0.0 to 5.0 mg P m^{-2} and day. 21 loading classes, in steps of 0.25 mg^{-2} and day were used for each of the 293 data vectors. Within these 21 loading classes, various ratios of particulate phosphorus to orthophosphorus, termed loading scenarios, were used. Scenarios of 0%, 10% and 25% orthophosphate in the total phosphorus loading were studied. Each scenario consisted of 21 * 293 = 6153 runs of the basic simulation model SEEMOD. For each loading class, 293 estimates of total primary production were thus obtained for each of the scenarios. The estimates have to be understood as the first year's reaction of the lake system to a certain loading, where the initial conditions at the beginning of that year reflect the empirical range of the lake's current state.

Primary production estimates were then arranged as a function of total phosphorus loading in a regression analysis. A linear and a power model were used, the linear model giving the larger correlation coefficient. An analysis of variance was performed to test the significance of the regressions ($F>592.$, d.f:20,N>6000;$P<0.001$). Regression curves for the three loading scenarios together with the 95% confidence limits for a new (single) estimate of primary production for a certain loading, are shown in Figure 6. Having the zero-loading class in common, the intercept does not differ significantly, whereas the slope of the regression curves increases with increasing

orthophosphate percentage in the total loading. Pooling the three scenarios, which seems justified as the actual orthophosphate ratio in the total phosphorus loading is uncertain, will therefore result in a progressively increasing uncertainty in the higher loading classes, as indicated in the divergent regression lines in Figure 6. For zero phosphorus loading, an average production of 45 gC m^{-2} and year is estimated, and the average production increase per mg P m^{-2} and day loading is predicted with 36 gC m^{-2} and year (r=0.8, N=6161). 95% confidence intervals for the slope estimate b are fairly narrow (below +/- 2.0%) and the differences between the scenarios are significant, indicating a significant correlation between orthophosphate percentage in the loading and the production per loading slope. However, the differences are small in absolute terms.

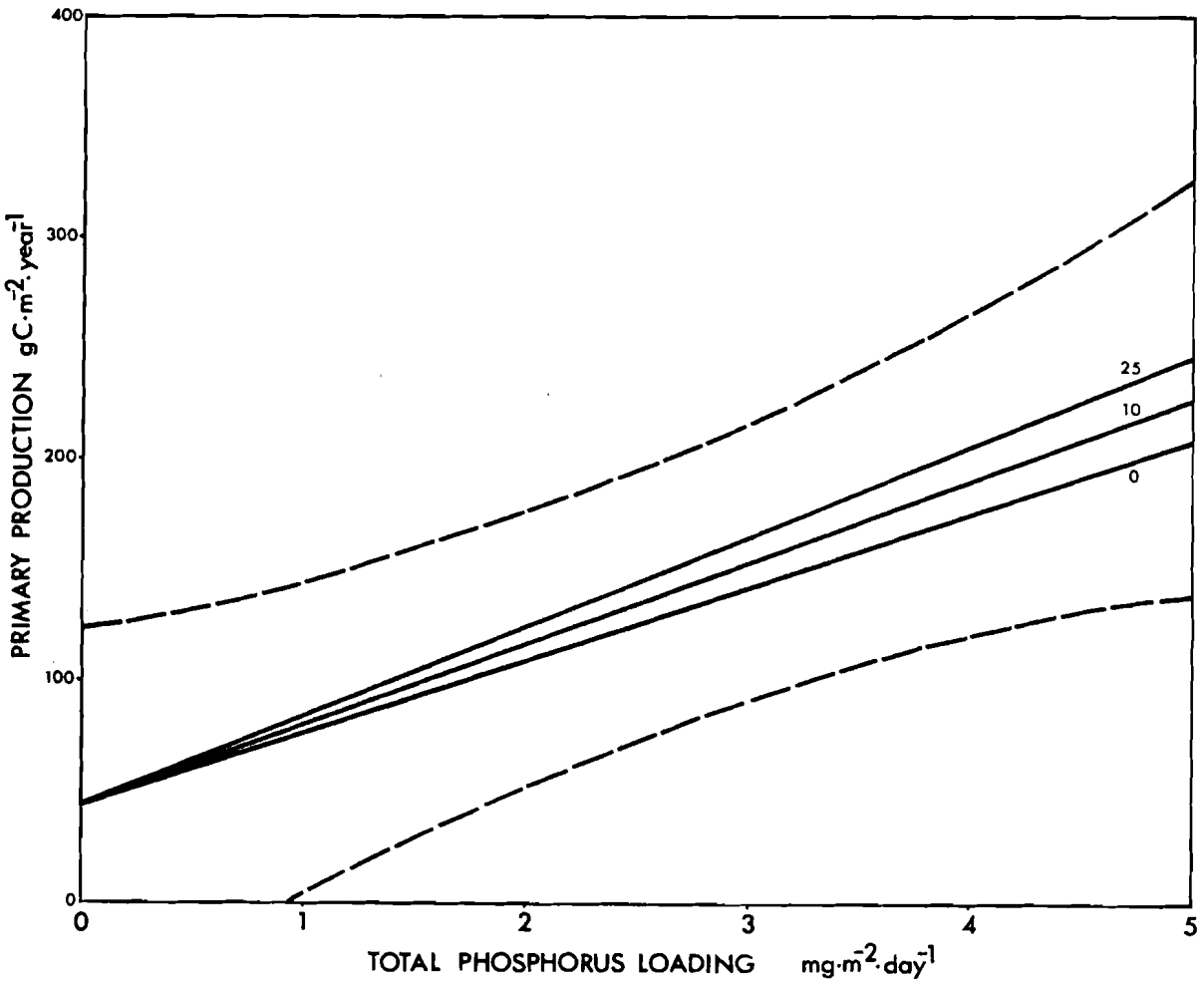

Figure 6. Loading/Production regression analysis for loading scenarios of 0%, 10%, and 25% solube reactive phosphorus in the total loading. 293 estimates per loading class and scenario. Broken lines indicate 95% confidence limits of production estimates for a given loading.

BEHAVIOUR SPACE AND OUTPUT PROBABILITY DISTRIBUTIONS

The estimates of states or outputs for a specific loading situation have to be regarded as random samples from a probabilistic behaviour space (for a discussion of this concept see pp. 38.39). Any specific measure can therefore be represented by the cross section of the behaviour space along one dimension, interpretable as a probability distribution. The programme ANALYSE3 was used to establish and analyse these distributions.

The estimates of primary production for each loading class were arranged as frequency distributions for production classes of 10 g C/m2 and year. As the estimates could be regarded as independent random samples from the behaviour space corresponding to the data space represented by the 293 "sample behaviour vectors" (see above), theoretical probability distributions were fitted for the sample estimates' frequency distributions. For an example see Figure 7. The area under these curves for a given range of production on the x-axis now represents the corresponding probability of the production to be within this range for a given "known" loading situation and the (specified) uncertainty in the residual knowledge about the system. Questions such as: what are the chances for the production to be below or above a certain level (or within a certain range) for a specific loading can now be answered (under the above assumptions on the residual uncertainty, the appropriateness of the chosen model, and the validity of the behaviour generating data sets outside the empirical range of behaviour, of course).

These probability distributions were now arranged in a three dimensional setup, where the production classes are shown on the x-axis, total phosphorus loading is represented in the z-axis, and the y-axis represents probability density. Figure 8 shows such an arrangement for the 10.0% orthophosphate loading scenario, which corresponds closely to the mean of field estimates of the phosphorus fraction proportions in the import to the Attersee.

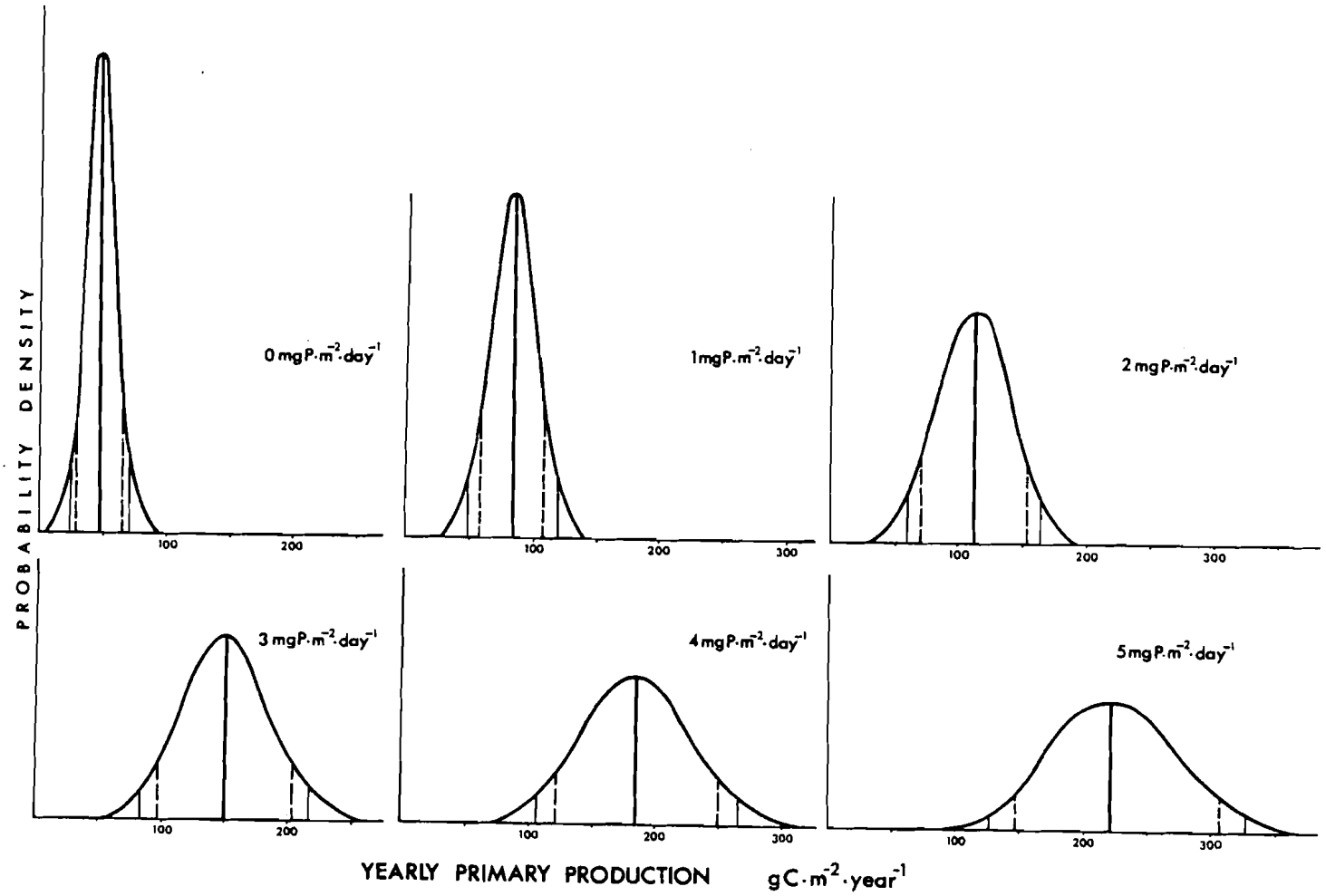

Figure 7. Loading/Production relationships. Curves show probability distributions fitted for the model output frequency distributions (class width: 10 gC m^{-2} and year). 90% and 95% confidence intervals for the mean estimates are indicated.

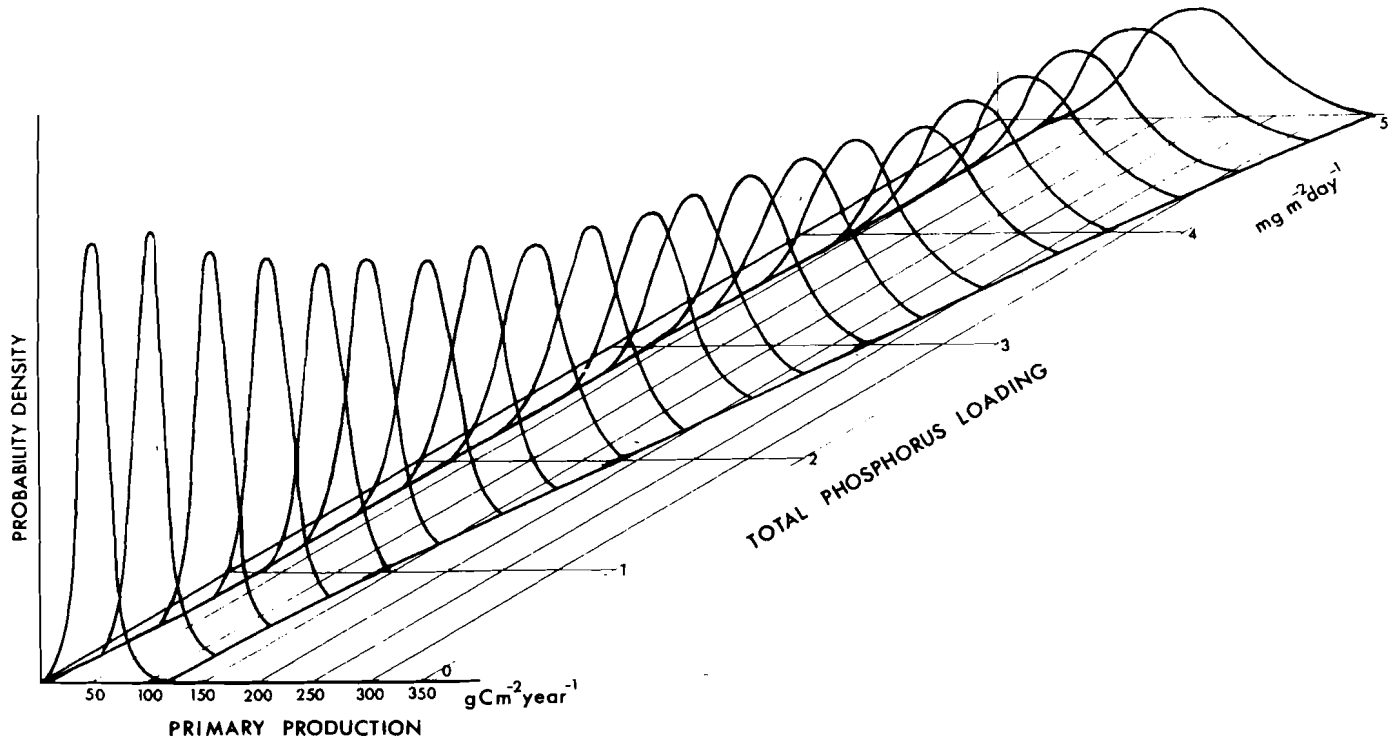

Figure 8. Loading/Production relationships for 21 phosphorus loading classes. Production is estimated for the first year of the indicated loading, starting with the empirical range of initial conditions; 10% loading scenario.

The perhaps most obvious feature of this representation is the increasing uncertainty in the high loading classes. This could also be seen as an increase in prediction uncertainty with increasing distance from the range of empirical inputs. To test this hypothesis, coefficients of variation (the standard deviation expressed as a percentage of the mean) were calculated

for the production estimates from the 21 loading classes. Plots of these coefficients vs. loading show a distinct minimum around 1 mg P m^{-2} and day, which corresponds to the mean loading of the behaviour runs (Table 3) and is close to the field estimate for average P loading (Figure 9). Therefore, as one could intuitively preclude, the precision of a prediction decreases with increasing changes in the inputs, or with increasing deviation from the descriptive case. In terms of the probabilistic behaviour space this would mean that the (normalized) volume is more and more distributed along its dimensions, showing no steep gradients along the axes.

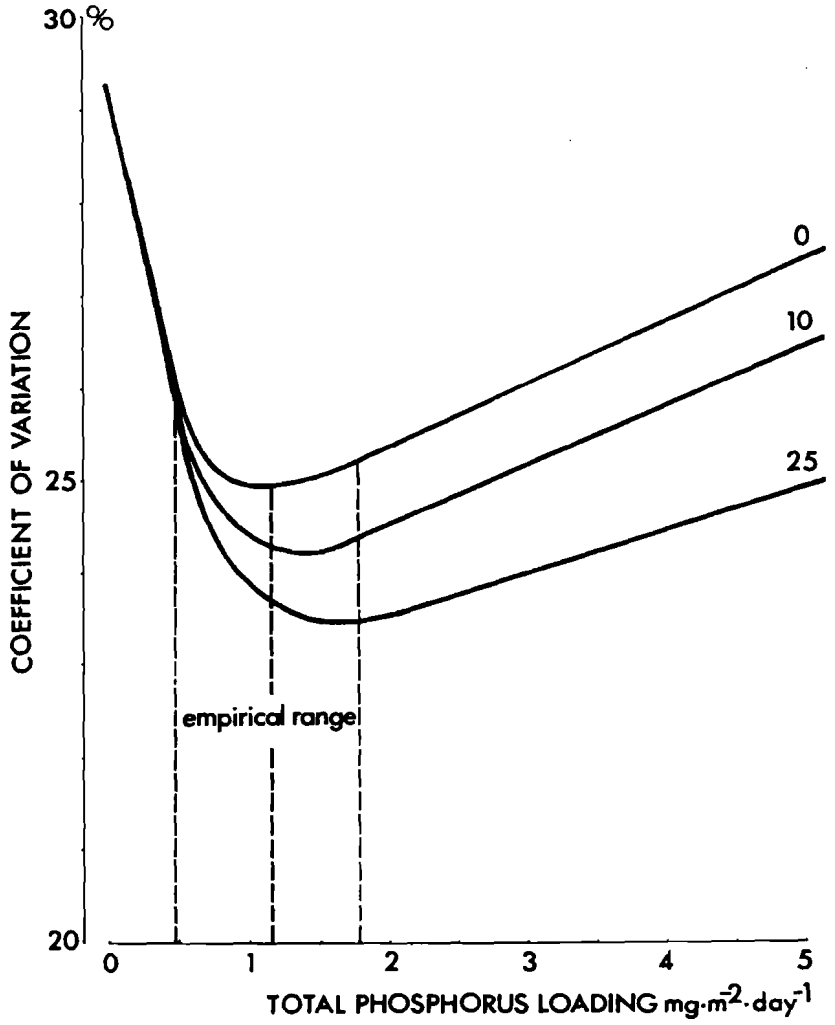

Figure 9. Coefficient of variation of production estimates for different P-loading.

LONG-TERM PREDICTION ACCURACY

The above series of estimates of yearly primary production per unit lake area for different phosphorus loadings predicts the lake's state in the first year after the change in the loading conditions. The initial states in this analysis reflect the empirical current (uncertain) state of the lake. The predicted production states can of course be of a transient nature, especially for larger changes in the loading conditions. The analysis was therefore extended for a ten years' period of changed loading conditions. Simulations were restricted to the 10% loading scenario. Starting with the range of initial conditions under the current loading, 150 vectors were used with the 21 loading classes for runs of ten years each. The state of the model was recorded for each year, using 16 behaviour describing data. Means, standard deviations, minima and maxima were then determined for each of these measures for each year and for all loading classes. The analysis of the long-term simulations (ANALYSE4) showed more or less unchanged and stable conditions (mean estimate from 150 runs each) in time for a total loading of 1.5 mg P m^{-2} and day. Below and above this level, the system experienced changes in its trophic state, returning to a new "equilibrium" after five to six years (Figure 10). Such hyperbolic patterns can be found in almost all behaviour measures investigated (comprising, besides yearly primary production, the minimum/maximum and start/end values of the state variables OP and PP, the day of the OP maximum, as well as yearly values for total P output, sedimentation, and net loading.)

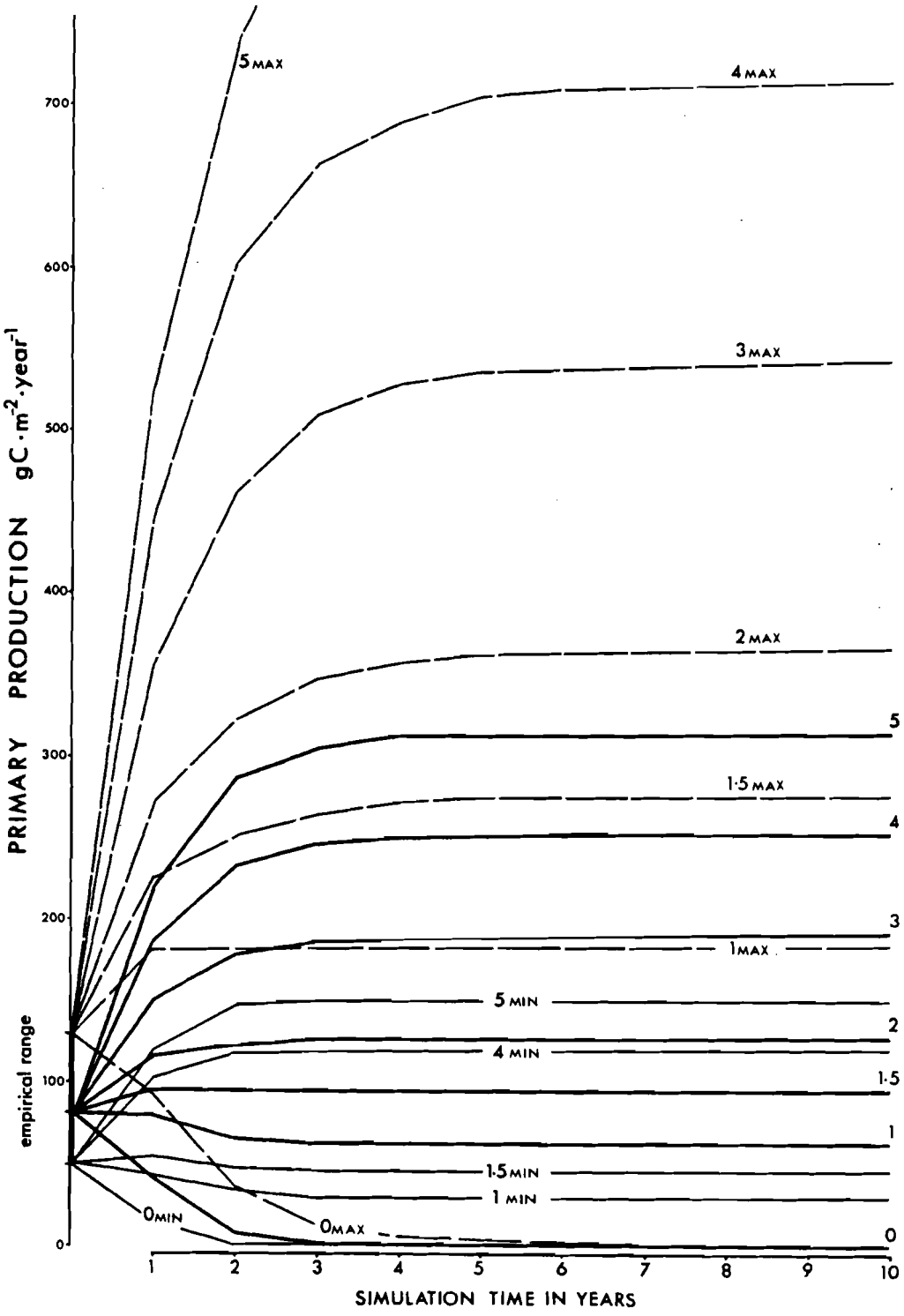

Figure 10. Mean estimates and ranges of yearly primary production from the long-term simulation runs for different total phosphorus loadings.

In most cases, the precision of the mean estimates (measured as coefficient of variation), shows a similar pattern in time, namely, rapid increase in the uncertainty in the first four to six years which then asymptotically approaches a certain level (Figure 11a). Also, the ranges of the estimates are diverging rapidly in the first few years and are then more or less constant. The relationship to the phosphorus loading classes or the distance from the empirical range of inputs is the same as described above (p. 30, Figure 9). However, some of the behaviour variables show quite different patterns. The coefficients of variation as well as the ranges of the estimates are decreasing in time, e.g., in the case of the yearly phosphorus sedimentation or the starting values of OP and PP, PP minimum, or the day of PP maximum (Figure 11b). The coefficient of variation as well as the ranges for the OP minimum is almost constant in time and is also independent from the loading class, being more or less constantly slightly below 100%. The OP maxima, in comparison, exhibit the "standard" behaviour with an asymptotic increase of the coefficient of variation and diverging ranges in time.

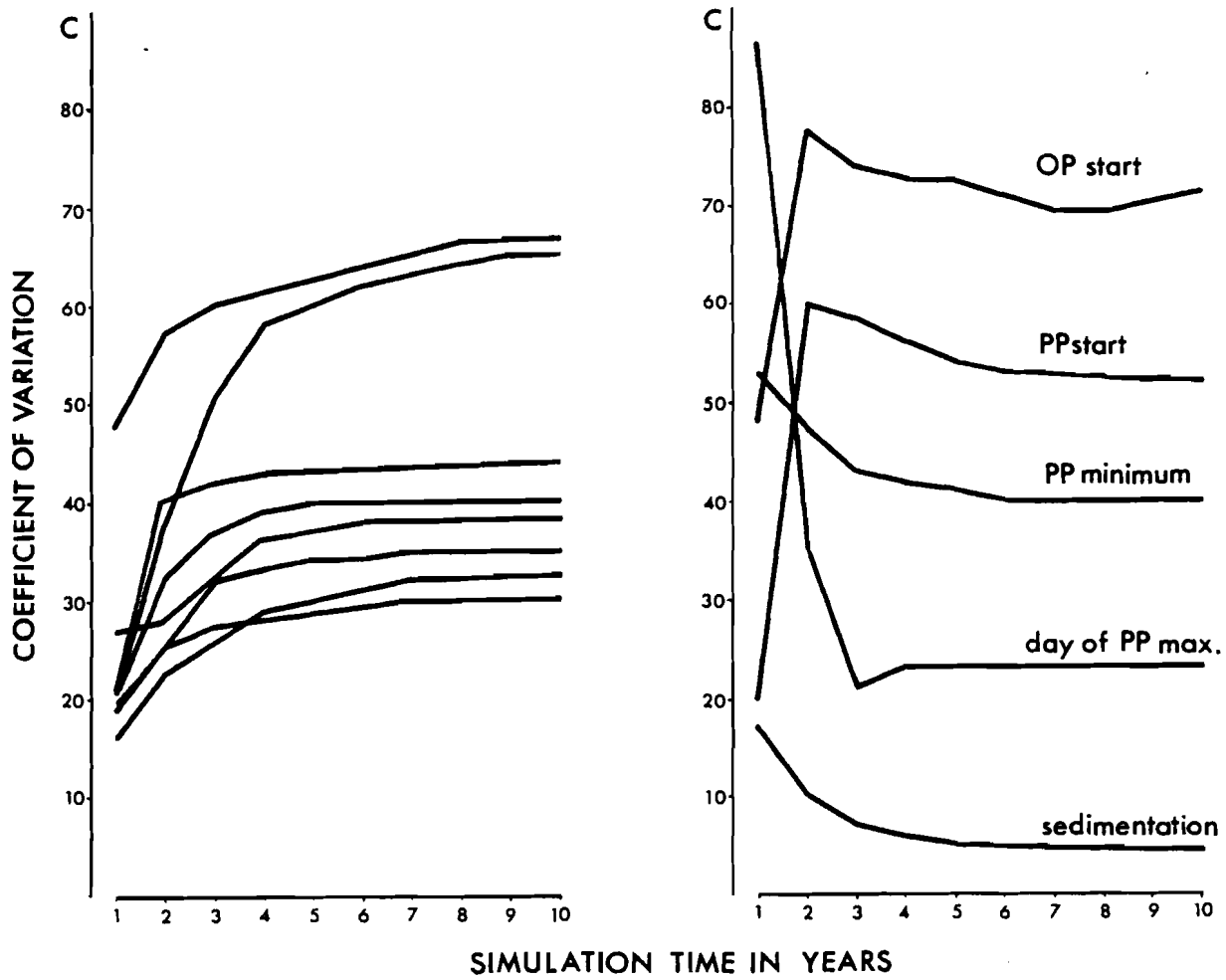

Figure 11. (a) Coefficient of variation in time for various behaviour variables and loading classes, showing the general pattern of asymptotic increase in time (left), (b) and some non-standard patterns (right).

DISCUSSION

The approach outlined above suggests the explicit inclusion of uncertainty in the data describing a system in the numerical modelling methods. To "preserve" some of the initial uncertainty throughout the numerical methods should allow to judge the accuracy and credibility of the final results of the modelling exercise. This does of course not reduce or at least make obvious the arbitrariness and uncertainty arising from other sources than the data required for the modelling--nevertheless, it is suggested as one step towards more rational modelling.

However, the methods proposed are rather brutal and straightforward. There are obvious technical limitations for this approach, especially when data spaces of high dimensionality are to be established. The only excuse for such an inelegant method is that the usefulness of complicated models of high state- and consequently data space dimensionality has to be questioned in principle, at least if predictions for highly complex and variable (and as a rule only badly known) systems are attempted. Not only the data have to be appropriate for a certain model, but also the model has to be appropriate for a certain problem with a certain system: that means rational in the sense of an earlier discussion (Fedra 1979). Only state variables and processes measurable in the field and relevant in terms of the problem addressed, and only model parameters which are directly physically interpretable should be used in such models, which consequently will be rather sparse.

The basic idea of the approach is the use of sets of "allowable" data and behaviour vectors or regions in data and behaviour space, respectively, instead of single, "optimum" (by arbitrary definition) vectors. The explicit variability of these sets represents the uncertainty in the information we have about the system, including statistical as well as stochastic variability.

In the language of set theory, which I used in the introduction for a compact notation, we can say, using the axiom of replacement (Fraenkel et al. 1973): if F is a function (the model) and D is a set, then there exists a set B which contains exactly the values F(d) for all members d of D which are in D'(F).

D'(F), the domain of F, is the class of all elements d for which there is a b such that $\langle d,b \rangle \in F$, and R(F), the range of F, is the class of all elements b for which there is a d such that $\langle d,b \rangle \in F$. Domain and Range of the model are here equivalent to corresponding regions in data- and behaviour-space respectively.

However, in the numerical approach we are dealing with random samples from the corresponding sets (domain and range) from the data and behaviour space respectively. Referring to Figure 1, the set BD, the population of all "realistic" behaviours (however unsatisfactorily defined from the ecological point of view) is well defined. But for the corresponding set DM in data space there is no operational definition at all. It is only (operationally) described by the sample DS'. Now, as

$$DS' \subset DD$$

and also

$$DS' \subset DM$$

it is obvious that the samples DS_i can only be drawn from the intersection of DM and DD:

$$DM' = DM \cap DD \quad .$$

Consequently, there exists the complementary subset DM''

$$DM'' = \{D_i \mid (D_i \in DM) \wedge (D_i \notin DD)\} \quad .$$

As the set DS* (which is used for the predictions) is basically derived from DS', we have to make the assumption

$$DM' >> DM'' \quad ,$$

which means that DS' can be taken as a representative sample from DM. DD represents our a priori knowledge about the parameters, forcings, imports, and initial states. If we specify

ample ranges on the basis of this knowledge for the DR_is (see pp. 6,15), we should expect that DM is included in DD and therefore

$$DM'' = \{\} \quad .$$

If, however, DM'' is of considerable size (which we can only estimate from the distributions of the d_i within the specified ranges DR_i, see p. 21 and Figure 4), we might take that as an indication that the model is structurally inadequate, or that the single data do not affect the model behaviour in correspondence to our conception of their physical and ecological (measured and estimated) counterparts. Consequently, the model could not be used for predictions: if "unrealistic" data vectors (DM'') give a reasonable behaviour within BD (representing our best knowledge of the empirical systems behaviours), the model as our theory on the processes and relations of the natural ecosystem is unable to reliably predict the systems reactions to changed "input" (in a broad meaning) conditions. Only if there is a consistent correspondence between data and behaviour space regions in terms of physical and biological plausibility and our empirical knowledge, the model could reasonably be used for predictions. The condition on DM'' is of course not independent from the definition of BD. The more rigid this behaviour definition can be made (without arbitrary and unjustified assumptions on field data accuracy, of course), the "sharper" the separation of DS into the two complementary subsets DS' and DS'' will be. Consequently, our knowledge about DM will increase.

A central issue of the proposed approach is the concept of a probabilistic behaviour space of systems and models. The concept of the behaviour space is somewhat different from the state space concept in the time domain. The state space can be defined as the set of all possible values which the state vector can assume at time t (Timothy and Bona 1968). In contrast, each of the behaviour space dimensions is defined for a

certain region (or discrete point) in time, which may be quite different for different dimensions. Measures with different time relations, such as daily means or yearly totals can be included, and the largest time interval over which a measure is determined, gives the order of the behaviour space. In practice and for models of ecosystems, this will generally be one year. In addition, these measures comprise properties of state variables and their relations as well as flows connecting them or outputs. The measures are primarily defined in accordance with the available data or measurable properties of a system (which of course are not restricted to the state variables of our models). Formally, we can write

$$B = \int_{t_o}^{t_e} F(t)\,[x(t_o);D(t)]\,dt \quad ,$$

where x is the state vector, t_o and t_e denote start and end-time of the observation or simulation cycle, and F is the set of algorithms used to derive the behaviour measures from the dynamic model.

However, for the description of a system, the meaning of this concept is quite obvious: for any interval in time there exists one point in a behaviour space (the maximum possible dimensionality of which corresponds to the time interval referred to) of a system, which describes its "state" in a general sense. And for any period in time containing several such intervals, there exists a set of behavioural events characterized by a mean behavioural vector and some variability around it or by the relative frequency (or probability) of behavioural events. However, our empirical knowledge is certainly incomplete and consists only of a few samples (including some measurement error). These samples are used to estimate the "true" behaviour means and variabilities. The description of a system in terms of an estimate of a mean behaviour vector only (if not in terms of a single (year's) data set, arbitrarily determined by availability), neglects part (and I believe a most important part) of the available information. For many ecological systems a mean behaviour (or "input") vector is of much less importance than the extreme values of some of its elements and the

probability of their coincidence. The description of such systems in terms of a probabilistic behaviour space seems therefore to be a useful concept, especially from the ecological point of view.

Considering now predictions from models, which are calibrated in a wide sense in reference to such a probabilistic behaviour description, it seems obvious that the probabilistic element has to be conserved. Its relative importance of course increases if we predict future behaviour, because to the uncertainty in the parameters and initial states there adds the uncertainty of future imports and forcings. The above example of course uses the simplest possible behaviour and also initial data space region structure by assuming rectangular probability density functions, and using the ranges only of the behaviour describing measures. However, with increasing knowledge about the system some of these data and measures could be described using centralized PDFs. The area of search for appropriate data vectors would consequently be restricted, and additional information on the expected distribution of the behaviour estimates, generated from random samples from the "appropriate" data space region, were available.

One of the merits of the above study may be that it includes all the data required by a model. On the other hand, this leads into some problems in the method for predictions, where some elements of the behaviour generating data vectors are varied systematically. The behaviour generating data space region of the model, explored under the constraints of the behaviour definition, is only roughly known by the points given by the behaviour vectors. Each of these vectors is of course a behaviour vector only for one value of all its elements, and when any of these elements is now varied systematically disregarding the correlation structure of the data space, additional uncertainty is introduced. In the above example, a significant correlation between particulate P loading d12 and the sedimentation- and diffusion-describing data d4, d5 and d7 was found in the data set GOOD2 (pp. 21,29). Therefore, as d12 was varied systematically, combinations of these data disregarding the correlation structure are likely to result in biased estimates. A selective use of the available data vectors according to the correlation matrix (Table 4) was considered.

Nevertheless, the results of the (technically much easier) unselective use of the vectors showed only a very small number of "outlayers" due to extreme data pairing, so that with regard to the high number of estimates no filtering procedure was applied to the results. However, this simplification will of course result in some overestimation of the scattering of the behaviour estimates. The increasing uncertainty of the estimates with increasing distance from the "empirical" range of the data-input (p. 30, Figure 9) can partly be attributed to this simplification in the numerical methods.

Another simplification in the model used, namely, setting to zero the optional backflux of phosphorus from the sediment, will also bias the results, especially in the very low and zero loading classes. Without doubt, the backflux of phosphorus from the sediments, although reasonably neglected in the empirical or higher ranges of phosphorus loading, will be of considerable relative importance in the case of very low phosphorus import. This is especially true for the long-term estimates of productivity, where a somewhat higher level of primary production could be expected. Generally, this leads back to the principle problem of a constant model structure and constant parameters, not related to inputs and states (p. 4). With regard to the above study, one might, for example, argue that the production estimates for the high phosphorus loading are unrealistic, because another nutrient such as nitrogen or silicon will become "limiting" when phosphorus levels are continuously increased. Again, uncertainty due to "structural uncertainty" also increases with increasing distance from the range of empirical conditions and behaviour.

If one is willing to accept all the partly rather arbitrary assumptions and simplifications on which this analysis is based, several conclusions for the lake system could be drawn. The first and perhaps most important one might be that the current knowledge about the lake system allows only very rough and rather trivial predictions of its future trophic state under changed loading conditions. The only way to improve the precision of such predictions will be to improve the data basis

for the analysis--and not to refine the models used. It is important to point out that improvement of the data basis requires model- or analysis-specific data. Only with a more confined initial data space and behaviour definition, the uncertainty of the predictions can be reduced. Consequently, only parameters which are directly physically interpretable should be used in ecosystem models designed for predictions, because only for such parameters the ranges can reasonably be confined on the basis of appropriate field data and experimental results. Another important requirement is that the level of aggregation in the model and in the measurements/experiments has to be at least balanced: monoculture laboratory experiments can hardly be used to estimate, e.g., a Michaelis constant for a model, which uses only one state variable for all primary producers. In situ experiments or appropriately designed laboratory experimentation have to account for high diversity and stochastic variability in ecological processes. Consequently, the number of measurements and experiments has to be increased and the spatial and timely scale has to chosen according to the respective scale of the models used. Physiological precision on the microscale simply is lost effort if the final analysis then represents the algae of several million cubic metres in a single number.

For the current example of the Attersee and the model SEEMOD2 this would mean that a specific sampling and experimental strategy had to be adopted in order to improve the data basis without increasing available manpower. Sampling could be restricted to two layers (above and below the thermocline) and one station only. At the same time, an increase in sampling units per sample (replicate sampling) should allow an estimate of the precision of sample means. Reducing the number of sampling dates during the mixed period would allow to increase sample density in spring, when rapid changes in phytoplankton and phosphorus concentrations are to be expected. Also, as the net sedimentation velocity of particulate phosphorus turned out to be of considerable importance for the model used, some additional experimental effort could help to improve the data basis required. Another example might be phosphorus import, where a restriction to total phosphorus only seems to be possible; on

the other hand, replicate sampling again would allow an estimate of the sampling error, which would help to define the data ranges for the model.

However, such suggestions are closely related to the model used and consequently to the specific purpose of the analysis. Different models with different underlying purpose might require different sampling and experimental strategies. Nevertheless, it seems important to point out the relationship between data collection strategies and the accessible precision of a problem specific analysis. Simple monitoring of a few standard variables, with unselective timing and spacing of samples, may well be insufficient for a specific analysis. Therefore, data collection and mathematical ecosystems analysis should be designed together in light of the environmental problems addressed.

SUMMARY

1. A stochastic approach to the quantification of model uncertainty is proposed, using a simple model of lake phosphorus dynamics. The approach suggests the explicit inclusion of data uncertainty in the numerical methods, and some straightforward methods of simulation and analysis are presented, which allow to describe a probabilistic model behaviour.
2. To define the behaviour of the lake system studied using uncertain data, ranges are specified for seven selected behaviour describing measures. Measures chosen are yearly primary production, yearly phosphorus export, allowable maxima for the state variables OP and PP, a time range for the algae spring biomass peak, and relational measures such as minimum relative increase of algae biomass during the year, or the maximum ratio of maximum to minimum total phosphorus in the lake during one year. The ranges for these behaviour variables can be understood as forming a box in a seven-dimensional behaviour space, representing in its extent the uncertainty in the environmental data as well as the stochastic variability of the ecosystem.
3. Twenty-two data-input values (comprising parameters, import/forcings, and initial conditions) are required for the simulation model used. Ranges are again specified for each of the data, and the resulting region in data space is randomly sampled using Monte Carlo techniques. More than 26,000 sample data vectors (from three slightly different data space

regions) are used for one-year simulations of the lake each, and the model behaviours obtained are classified according to the above behaviour definition.

4. Behaviour generating sample data vectors are now used for another series of almost 50,000 simulation runs, where the P-loading determining data are systematically varied in 21 classes from 0.0 to 5.0 mg/m2 and day. Several ratios of OP:PP in the import are studied, and simulations are extended for up to ten years.
5. The estimates of behaviour variables obtained in these simulations are analysed, and probability distributions are fitted for the relative frequencies of the behaviour variables for different P-loading situations.
6. The analysis of the propobabilistic model behaviour indicates clearly that prediction uncertainty increases with increasing distance from the empirical range of inputs. Prediction uncertainty also increases with simulation time, asymtotically approaching a certain level, which depends on the initial data uncertainty (reflected in the variability of the sample data vectors used) as well as on the distance from the empirical range of input conditions/behaviour.

REFERENCES

Attersee (1976) Vorlaeufige Ergebnisse des OECD- Seeneutrophierungs- und des MaB-Programms. 197 pp.

Attersee (1977) Vorlaeufige Ergebnisse des OECD- Seeneutrophierungs- und des MaB-Programms. 148 pp.

Beck, M.B., E. Halfon, and G. van Straten (1979) The Propagation of Errors and Uncertainty in Forecasting Water Quality. Laxenburg, Austria: International Institute for Applied Systems Analysis.

Di Toro, D.M., and G. van Straten. Uncertainty in the Parameters and Predictions of Phytoplankton Models. WP-79-27. Laxenburg, Austria: International Institute for Applied Systems Analysis.

Fedra, K. (1979) Modelling Biological Processes in the Aquatic Environment (with special reference to adaptation). WP-79-20. Laxenburg, Austria: International Institute for Applied Systems Analysis.

Floegl, H. (1974) Die Reinhaltung der Salzkammergutseen. Oesterr. Wasserwirtschaft 26,1/2, 5-11.

Fraenkel, A.A., Y. Bar-Hillel, and A. Levy (1973) Foundations of Set Theory. 404 pp. Amsterdam: North-Holland Publishing Company.

Imboden, D. and R. Gaechter (1978) A Dynamic Lake Model for Trophic State Prediction: Ecol. Model. 4, 77-98.

Lewis, S., and A. Nir (1978) A Study of Parameter Estimation Procedures of a Model of Lake Phosphorus Dynamics. Ecol. Model. 4, 99-118.

Patten, B.C. (1971) Systems Analysis and Simulation in Ecology I. New York: Academic Press.

Spear, R.C., and G.M. Hornberger (1978) Eutrophication in Peel Inlet: An Analysis of Behaviour and Sensitivity of a Poorly Defined System. CRES Rept. AS/P24.

Straskraba, M. (1976) Development of an Analytical Phytoplankton Model with Parameters Empirically Related to Dominant Controlling Variables. Symposium fuer Umweltbiophysik, Abh. Akad. Wiss. DDR Jg. 1974 Adademieverlag Berlin 1976.

Timothy, L.K. and B.E. Bona (1968) State Space Analysis: An Introduction. 406 pp. New York: McGraw-Hill.